



# Özbekçe↔Türkçe Otomatik Çeviri Yazılımı İçin Deyimlerin Veritabanını Oluşturma Sorunu

Manzura ABJALOVA<sup>1</sup>

Umida RAŞIDOVA<sup>2</sup>

## Öz

Deyimlerin, birkaç kelimeden oluşan ortak bir dil birimi ve harfi harfine tercüme edilemeyen sözdizimsel bir bütün olarak kabul edildiği bilinmektedir. Linguo-kültürel öneme sahip deyimler, her milletin tarihi boyunca hayatındaki çeşitli durumların bir sonucu olarak sembolik benzetmeler ve karşılaştırmalar sonucunda oluşmuştur. Çevirideki metne linguo-kültürel yaklaşım, sanatsal söylemin kültürel olarak tanımlanmış idio-etnik, kültürel farklılığının neden olduğu engellerin belirlenmesine ve aşılmasına olanak tanır. Bu nedenle, otomatik çeviride verimlilik elde etmek için, deyim birimlerinin (DB) paralel bir veritabanını oluşturmak önemlidir. Bu makale, deyimlerin sözcüksel bir birim olarak tanımlanmasını, deyimlerin temelini oluşturulmasını sağlayan faktörleri incelemektedir.

**Anahtar kelimeler:** Deyimsel birim (DB), deyim, kelime, cümle, otomatik çeviri, dilsel veritabanı.

## Problem of Creating a Database of Idioms for Uzbeki-Turkish Automatic Translation Software

### Abstract

It is known that in today's globalized era, language corpora are a kind of teaching tool, a multifunctional linguistic electronic system. The presence of lexicographic resources in it along with the text base as a source of information gives users the opportunity to acquire basic knowledge and replenish their vocabulary. For this purpose, it is considered important to have a phraseme database in the lexicographic database. This article discusses the reasons for ensuring the existence of a database of expressions in linguistic corpora and its meaning. The methods of working with phrases in the Uzbek language educational corpus were also shown.

<sup>1</sup> Assoc. Prof. Dr., Tashkent State University of Uzbek Language and Literature. Tashkent, Uzbekistan. E-mail: abjalova.manzura@gmail.com

ORCID: 0000-0002-1927-2669

<sup>2</sup> Assoc. Prof. Dr., Samarkand State University, Uzbekistan. E-mail: umida\_rashidova@bk.ru



**Keywords:** Phrases, language corpora, educational corpus, base, somatic phrases.

## GİRİŞ

Özbekçe ve Türkçe tipolojik olarak Altay dil ailesine aittir ve akraba diller olarak kabul edilir. Geneolojik ve tipolojik olarak, bu ilgili diller, özellikle sözcüksel, morfolojik ve sözdizimsel düzeylerde benzerliklere sahiptir. Her iki dildeki yüksek derecede tipolojik benzerlik ve evrenselliğe rağmen küresel ağda yapay zeka yöntemlerine dayalı olarak “Google translate” ve “Yandex translate” hizmetleri tarafından da metinlerin Özbekçeden Türkçeye ve tersi, Türkçeden Özbekçeye çevrildiğinde morfolojik ve sözcüksel hatalar, üslup hataları olabilir. Mesela, Özbekçe *Bu ishda sizning qo'lingiz bormi?* 2. çoğul şahıs (siz zamiri çoğul değil, saygı anlamında)ki cümle Türkçeye “Google Çeviri”de *Bunda parmağın var mı?* 2. tekil şahıs olarak verilmiştir. Başka bir örnek, *Men siz bilan yelkama-yelka ishlaganimdan xursandman.* cümlesi “Google Çeviri”de Türkçeye “*Seninle çalışmaktan mutluyum*” tarzında ve “Yandex Çeviri”de “*Seninle omuz omuza çalıştığım için mutluyum*” tarzında çevrilmiştir. Her iki sistemde de ortak olan hata, saygının anlamının korunmaması ve özgünlüğün “Yandex Çeviri”de “*omuz omuza çalışmak*” deyiminin Türkçe alternatifine çevrilmesiyle belirlenmesidir. Bu deyim “Google Translate” hizmetinde Türkçe olarak kaydedilmemektedir. Ayrıca *-dan* uzaklaşma hâli eki Türkçede *-dan* şeklinde değil, *için* son çekim edadı kullanılarak çevrilir. Bu tür morfolojik, sözlüksel ve deyimsel eksiklikler, Özbekçe ↔ Türkçe için bir çeviri programının kusursuz dil desteği temelinde oluşturulmasını gerektirmektedir. Kuşkusuz, kullanılan tipolojik veritabanı, bağımsız parametrelerin mantıksal olarak yaygın evrensellik ve farklılaşma özelliklerine göre geliştirildiğinde daha da önem kazanmaktadır.

Çeviribilimde çeviri öncesi metin incelemesi önemlidir. Metnin dilsel birimleri (ses, biçimbirim, kelime, tamlama ve cümle) çeviriden önce dilbilimsel incelemeye tabi tutulur.

## KAYNAK ANALİZİ

Çeviri öncesi dilbilimsel analizde, anlamlı parçalar çeviri birimleri olarak kabul edilir ve bunlar fonetik, sözcüksel, dilbilgisel (morfolojik, sözdizimsel) ve stilistik olarak analiz edilir. Makine çevirisi için kelime kombinasyonlarının, deyimsel kombinasyonların, deyimlerin, paremlerin (aforizmaların ve atasözlerin) çevirisi konusu çok acil bir sorundur. Bu kadar karmaşık yapıya sahip ancak ortak anlambilime sahip birimlerin alternatif çevirisini gerçekleştirmek için çeviri programlarına paralel bir veritabanını oluşturmak önemlidir. Bu durumda sözcük çözümlemesine ait deyimsel birimlerin dikkate alınması uygundur.

## DEYİMSEL BİRİMLER KELİME / CÜMLEYE EŞİTTİR (DEYİM = KELİME / CÜMLE)

Sebepler:

A. Anlam uyumu, anlamsal bütünlük.

Birincisi, DB'deki kelimeler anlamsal olarak tutarlıdır ve anlamlı bir bütünlük oluşturur. Çoğu zaman, anlam tek kelimeyle aynıdır ve taşınabilir. Semantiğin tam olarak bir kelimeye eşdeğer olması, deyimleri kelimelerle eşitlemenin de temelini oluşturur.

B. Deyimler dilde hazır birimlerdir.

İkincisi, bir deyim, bir tamlama veya bir cümle gibi konuşma sürecinde oluşmaz, tıpkı bir kelime gibi dilde hazırdir. Yapısal olarak, bir tamlamaya veya bir cümleye eşdeğerdir, ancak anlamı bir kelimeye eşdeğerdir. Bu nedenle, deyimlere bir “kelime tanımlayıcısı” (Балли, 2001) ve bir “cümle tanımlayıcısı” (Абжалова, 2020) vermek uygundur. Bileşikleri kimyasal bileşiklerle karşılaştıran Ş.Bally, “Deyimsel birim daha yaygınsa, o zaman açıkçası, bu durumda, deyim basit bir kelimeye eşdeğerdir” (Балли, 2001) şeklinde inanıyordu.

B1. DB=kelime indikatörü

*bosh og'rig'i / baş belası = dahmaza* (ortiqcha tashvish / aşırı kaygı)

*kapalagi uchmoq / korktuğuna uğramak = qo'rqmoq / korkmak*

B2. DB=cümle indikatörü

*Holiga maymunlar yig'laydi / Haline maymunlar ağlıyor (Haline köpekler gülmüş) = Ahvoli yomon. / Durum kötü.*

*Devorning ham qulog'i bor / Yerin de kulağı var (Duvarın da kulağı var) = Kimdir eshitib qolishi mumkin. / Birileri duyabilir.*

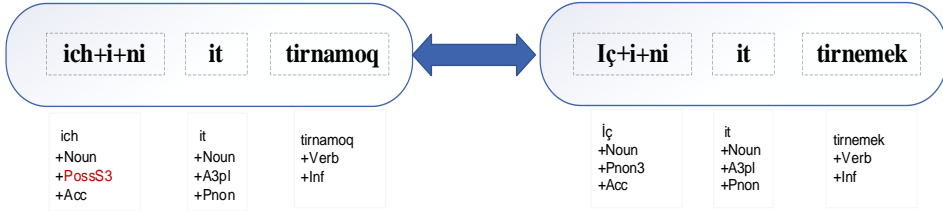
DB'lerdeki paradigmatik ve dizimsel ilişki, metin işleme, otomatik düzenleme ve metin analizi ve makine çevirisi yapan programların dil desteğinde özel bir alan oluşturduğunu göstermektedir (Abjalova ve Rashidova, 2021).

**Tablo 1. Cümle Yapısı Deyimin Çekimi**

Bir tür cümle yapısı deyimin çekimi			
Tekil	Modeli	Çoğul	Modeli
<i>ko'zim pinakka ketdi</i> (uykuya daldım)	$N_{e.a.1b} + N_{k.a} + V_{1b}$	<i>ko'zimiz pinakka ketdi</i>	$N_{e.a.1k} + N_{k.a} + V_{1b}$
<i>ko'zing pinakka ketdi</i>	$N_{e.a.11b} + N_{k.a} + V_{1b}$	<i>ko'zingiz pinakka ketdi</i>	$N_{e.a.11k} + N_{k.a} + V_{1b}$
<i>ko'zi pinakka ketdi</i>	$N_{e.a.111b} + N_{k.a} + V_{1b}$	<i>ko'zlari pinakka ketdi</i>	$N_{e.a.111k} + N_{k.a} + V_{1b}$

C. Deyimlerin bire bir değil, bütünsel bir çevirisi vardır.

Üçüncüsü, bir deyimın bileşenleri sabit bir kalıpta sabitlendiğinden, bir dilden diğere tam anlamıyla çevrilemez. Birebir çeviri, düşünce karmaşasına, çeviri çılgınlığına ve hatta anlam kaybına yol açar. Bu nedenle otomatik çeviri programları, mobil uygulamalar, platformlar hatta paralel derlemlerin dil desteği oluşturulurken, tamlama ve cümle yapısına sahip deyimler sözcüksel birimler olarak veritabanında yer almakta ve diğere dillerde alternatif çeviriler de sağlanmaktadır. Örneğin, *ichini it tırnamoq* deyimini “Google Çeviri”de “köpeğın içinı tırmalamak” ve “Yandex Çeviri”de “köpek tırmalamak içinde” olarak belirsiz çevrilmiştir. Aslında Türkçede bu deyimın içinı it tırnemek gibi alternatif bir deyimı var.



**Çizelge 1. "Ichini it tırnamoq" (Özbekçe) / "İçini it tırnemek" (Türkçe) alternatif deyimlerinin morfolojik olarak eşitlemesi.**

Cümle yapısında 3. tekil şahıs iyelik eki (PossS3) kırmızı renkle verilmiştir ki bu da deyimın burada iyelik kategorisine göre tasnif edildiği anlamına gelmektedir. Deyimleri modellerken bu hususa dikkat etmek önemlidir. Özbekçe ve Türkçe deyimlerin çekimi (Tablo 1) ve fiil çekimi yerini paralel veritabanda gösteren modellerin çeviri doğruluğunu artırmada önemli olduğu düşünülmektedir. Bu durumda DB=kelime/cümle kavramı önem kazanmaktadır (Abjalova ve Abdumuminov, 2021).

Tipolojik olarak Özbekçe ve Türkçe, Altay dil ailesinin Türk dil grubuna aittir ve bu tür sözcüksel ve gramer doğruluğu oluşturur. Bu tür bire bir alternatif deyimler azdır, hatta bazı bire bir deyimler farklı bir anlam ifade eder. Örneğin;

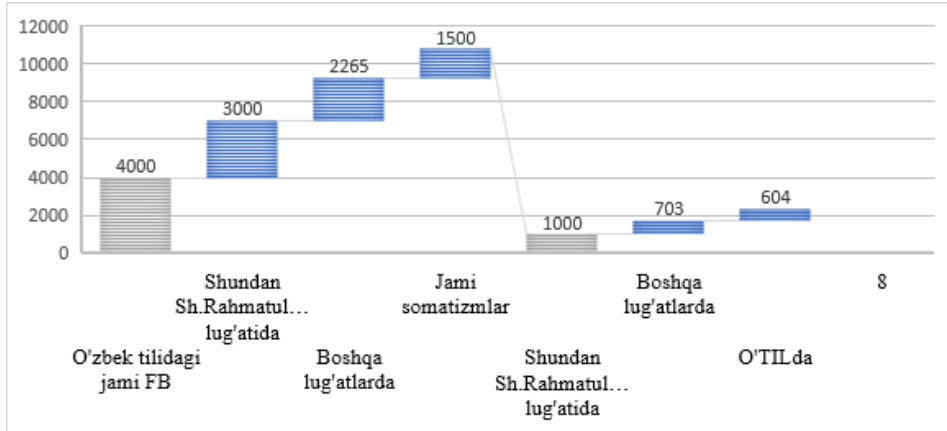
Türkçe göğüs *germek* deyimı hayatın bütün zorluklarına katlanmak, *dişlerini dişlerine koymak*, hayatın bütün engellerine ve darbelerine rağmen geri adım atmadan ilerlemek, hedefe yürümek demektir (Amanova, 2014). Özbekçe karşılığı olan “ko‘krak kermoq” deyimının anlamı tamamen farklıdır. “Ko‘krak kermoq” deyimı, belli bir işi iyi yaptığınızda, başarılı olduğunuzda gurur ve kibir ifade etmek için kullanılabilir. Ayrıca Özbek dilinde bu deyim genellikle başarılı, iyi bir oğulun ebeveynleri için kullanılır. Türkçedeki göğüs *germek* ile Özbekçedeki *ko‘krak kermoq* deyimleri sözlük ve yapı olarak aynı olsa da kullanım yerleri farklıdır. Bu iki deyimın temel anlamı anlaşılmadan

## Özbekçe↔Türkçe Otomatik Çeviri Yazılımı İçin Deyimlerin Veritabanını Oluşturma Sorunu

Türkçeden Özbekçeye bire bir tercüme edilirse “göğüsü germek”, “göğüsün gerilmesi” ve “göğüsle yürümek” gibi çok saçma çeviriler ortaya çıkıyor.

Aynen çeviri öncesi analizi sürecinde dil dışı faktörleri hesaba katmak önemlidir. Örneğin, deyimlerin çözümlenmesinde linguo-kültürel ve linguo-pragmatik etkilerin dikkate alınması sonucunda ulusal rengin korunması sağlanmakta, çevirinin yeterlilik ve alternatiflik düzeyi artmaktadır.

Özbekçe ifadeler Ş.Rahmatullayev (Rahmatullayev, 2001), Özbek Dilinin Açıklamalı Sözlüğü – O’TİL (O’zbekiston milliy ensiklopediyasi, 2006) ve diğer sözlükler aşağıdaki sayıları içerir: Özbekçe toplam DB – yaklaşık 4000: bunlardan Ş.Rahmatullayev’in sözlüğünde yaklaşık 3000; diğer sözlüklerde 2265’tir. Somatik deyimlerin toplam sayısı yaklaşık 1.500’dür: bundan Ş.Rahmatullayev’in sözlüğünde yaklaşık 1.000 tane bulunur; diğer sözlüklerde 703, O’TİL’de ise 604’tür (Çizelge 2). Türkçedeki toplam deyim sayısı 10.000 civarındadır (Turkcha lug’at).

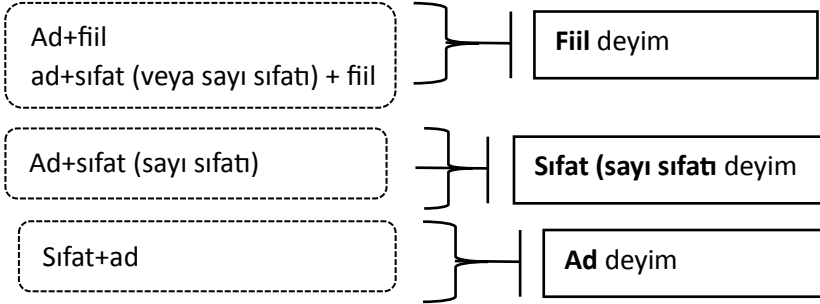


Çizelge 2. Özbek dilinde deyim ve somatizm istatistikleri

### SOMATİK DEYİMLER

Her dilde insan vücudunun bölümlerini içeren somatik deyimler vardır ve bunlar artık Özbek dilbiliminde özel bir araştırma konusudur (Rashidova, 2018). Özbekçe somatik deyimlerin sayısı binden fazladır. “Özbek Dilinin Açıklamalı Sözlüğü”nde *göz bileşenli* deyimlerin sayısı 80’den fazla iken, Ş.Rahmatullayev’in “Özbek Dilinin Deyimsel Sözlüğü”nde 158 olduğu belirtilmektedir. A.İsayev’in “Özbek Dilinde Somatik Deyimler” başlıklı araştırması *göz bileşenli* deyimlerin Özbekçede 127 tane olduğunu bildirmiştir. Araştırmacı, *baş*, *göz*, *el* ve *ayak* kelimelerinin bulunduğu DB sayısının Özbekçede 770, Türkçede 675, kombinasyona eşit olan deyim sayısının ise Özbekçede 510, Türkçede 495 olduğunu belirlemiştir (İsayev).

Cümleler, insanın sınırsız ve çeşitli duygularını ve fizyolojik süreçlerini çekici ve sanatsal bir şekilde ifade etmeye hizmet eder. Özbekçe somatik deyimlerin ana kısmı aşağıdaki yapısal modellere sahiptir (Çizelge 3):



Çizelge 3. Somatik deyimlerin gramer modelleri.

Örneğin, *ko'z yugurtirmoq*, *ko'ziga issiq ko'rinmoq*, *ko'ziga yaqin ko'rinmoq*, *ko'zi ola-kula bo'lmoq*, *ko'zi to'rt bo'lmoq*, *ko'zi ochiq ketmoq*, *ko'zi ochiq*, *ko'zi bejo*, *ko'zi taka-puka* gibi.

**Deyimleri modelleme.** Tüm bu DB'ler, kişinin belirli bir içsel zihinsel durumunu ifade etmeye hizmet ettiği için konuşmacı, dinleyici veya başka bir kişi olmak üzere üç kişiden birine aittir. Bu nedenle, DB'de isim grubuna ait bir kelime, çekimli bir kelime formu veya fiil grubuna ait bir kelime gramer göstergeleri alır. Mesela, *Qisqasi, endi unga xat yozishga qo'lim bormadi* (A.Muxtor). Türkçe çeviri: *Kısaca artık ona bir mektup yazmaya elim gitmedi*. *Qo'lim bormadi* ← *qo'li bormadi* ← özne+yüklem ilişkisini oluşturan bir DB formu: *elim* (-m iyelik kategorisinin birinci tekil şahıs eki=>Ne.a.I.b + V<sup>x</sup>), yani <*elim gitmedi*> → <*elim*> -im => Ne.a.I.b + V<sup>x</sup>.

Bu savoldan uning kayfi tarqalgandek bo'lib, *ko'zlari* (-lar çoğul eki, -i iyelik kategorisinin üçüncü tekil şahıs eki=>Ne.a.III.k.) *moshdek ochildi* (-il dönüşlü görünüş göstergesi, -di şahıs kategorisi III tekil kişi göstergesi=>Avd-Vpass.v.+III.b.) (Oybek. Çocukluk),

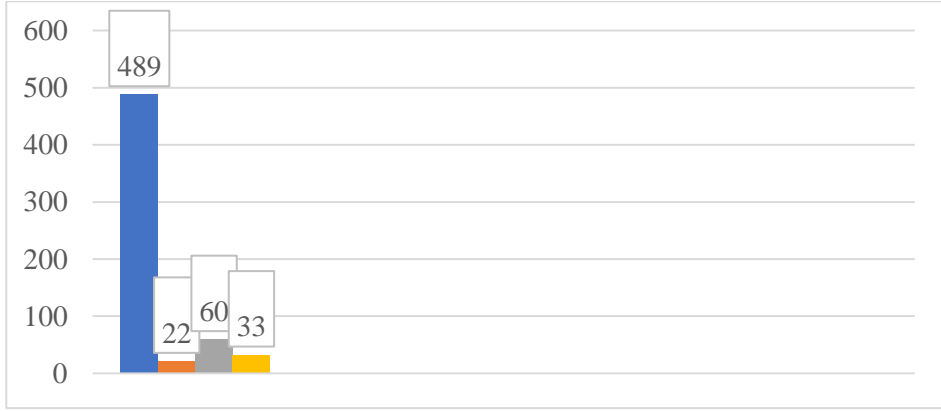
Türkçe çeviri: *Ruh hali bu sorudan dağılmış gibiydi ve gözleri iyice açıldı*.

"Qani, Muhiddin aka, *yuragingiz ham kuygandir*, qovunga marhamat", - dedi Muhiddin (R.Fayziy. Çöle ilkbahar geldi).

Türkçe çeviri: "Hadi Muhyiddin ağabey, *belki yüreğin yanıyor*, buyur kavuna" dedi.

*Yuragi kuymoq* FB'nin gramer yapısı: *yuragingiz* – lemma + (-ingiz iyelik kategorisinin 2. çoğul şahıs eki=>Ne.a.II.b.) *ham* – sona gelen edat, *kuygandir* – lemma + geçmiş zaman (-gan) + sona gelen edat (-dir).

## Özbekçe↔Türkçe Otomatik Çeviri Yazılımı İçin Deyimlerin Veritabanını Oluşturma Sorunu



**Çizelge 4.** Ş. Rahmatullayev'in sözlüğündeki somatik deyimlerin kategorilere göre istatistikleri

Ş. Rahmatullayev'in sözlüğündeki somatik deyimlerin toplam sayısı 604'tür, kategorilere göre şu istatistikler vardır: fiil – 489, isim – 22, sıfat – 60, zarf – 33 (Çizelge 4).

### SONUÇ

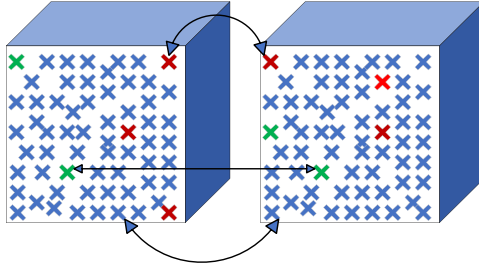
Özbekçe ve Türkçedeki deyimlerin analizi sonucunda şu sonuçlara ulaşılmıştır (Çizelge 5):

1. Bire bir uygun. Tipolojik olarak akraba Özbekçe ↔ Türkçedeki deyimler tam anlamıyla tercüme edilebilir. Mesela, *eshikdan quvsang teshikdan kiradi – kapıdan kovsan bacadan girer, joni halqumiga kelmoq – can boğazdan gelmek, ko'z quloq bo'lmoq – göz kulak olmak, tindirmaslik – nefes aldirmamak, jonini bermoq – canını bağışlamak*. Ancak bu uyum küçük bir sayıdır (Çizelge 5'te yeşil renkte).

2. Bire bir uygun, ancak pragmatik uygulama farklıdır. Kelimenin tam anlamıyla tercüme edilen deyimler, her iki halkın temsilcileri tarafından özel durumlarda kullanılabilir. Elbette *linguo-kültürel* olguyla bağlantılıdır. Örneğin, *göğüs germek* deyiminin Özbekçeye bire bir tercümesi “*ko'krak kermoq*” deyimini olmasına rağmen, kullanım durumuna göre *tishini-tishiga qo'ymoq* deyimini ile pragmatik olarak tutarlıdır (Çizelge 5'te kırmızı renkte).

3. Anlamsal olarak tutarlı olan ancak sözcüksel birimleri tutarsız olan deyimler. Birbirinden tamamen farklı sözcük birimlerinden oluşturulmuş deyimler vardır, ancak içerik ve kullanım durumları birbiriyle uyumludur, çünkü bu tür deyimler Özbek ve Türk halklarının yaşam ve kültüründe belirli durumlar, teşbihler ve karşılaştırmalar sonucunda ortaya çıkmıştır. Mesela, *tovbasiga tayanmoq – dizini dövmek* (juda pushaymon bo'lmoq), *ikki oyoqni bir etikka tiqmoq – yumurta kapıya dayanmak* (bir ishga shoshilayapti, ammo vaqt kam

qolganida qo'llaniladi), o'ziga choh qazimoq – bindiği dalı kesmek, boshi ko'kka yetmoq – etekleri zil çalmak (o'ta quvonmoq) (Çizelge 5'te mavi renkte).



**Çizelge 5.** Özbekçe ve Türkçedeki deyimlerin uyumluluğunu gösteren resim.

Deyimlerin bilgisayar modellemesi ve paralel veritabanlarının oluşturulması aşağıdaki olasılıkları sağlar (Abjalova ve Rashidova, 2023):

- 1) deyimsel birimlerin otomatik analizini kolaylaştırır;
- 2) metindeki deyimsel birimlerin çevirisini kolaylaştırır;
- 3) dilsel işlemcide bir kelime veya cümle tanımlayıcısı olarak anlamsal olarak anlaşılır;
- 4) linguo-kültürel önem korunur;
- 5) deyimlerin anlamsal bütünlüğü korunur.

Bu makalede, deyimlerdeki sözcüklerin sabit birleşimleri, deyimlerin anlamsal bütünlüğü ve linguo-kültürel önemi, serbest sözdizimsel bağlantısı incelenmiştir. Bu özelliklerden dolayı, bu tür kararlı bileşiklerden paralel bir veritabanı oluşturmanın önemi örneklerle kanıtlanmıştır. Çeviride deyimlerin kelime/cümle olarak değerlendirilmesinin nedenleri, deyim=kelime/cümle kavramı açıklanmıştır. Genel olarak, DB'lerin paradigmatik ve dizimsel oluşum süreci sosyal önem kazanır ve karmaşık bir yapısal süreç olarak kabul edilir, özellikleri metin ve konuşma durumunda daha net bir şekilde ortaya çıkar. Bununla birlikte deyimlerin bir tamlama ve cümlenin yapısal birimi, hazır sözcük birimi olarak ele alınması ve sözdizimsel ilişkiler biçimindeki yerinin belirlenmesi uygun görülmüştür.

## ETİK BEYAN

“Özbekçe↔Türkçe Otomatik Çeviri Yazılımı İçin Deyimlerin Veritabanını Oluşturma Sorunu” başlıklı çalışmanın yazım sürecinde bilimsel kurallara, etik ve alıntı kurallarına uyulmuş; toplanan veriler üzerinde herhangi bir tahrifat yapılmamış ve bu çalışma herhangi başka bir akademik yayın ortamına değerlendirme için gönderilmemiştir.



## ÇATIŞMA BEYANI

Çalışmada herhangi bir potansiyel çıkar çatışması söz konusu değildir.

## KAYNAKÇA

- [1] Балли Ш. Французская стилистика / Ш. Балли. – М.: Эдиториал УРСС, 2001. – 392 с. (Bally Sh. French style / S. Bally. - M.: Editorial URSS, 2001. - 392 p.)
- [2] Абжалова, Манзура (2020). Автоматический анализ фразеологических единиц в лингвистических программах // VIII Международная конференция по компьютерной обработке тюркских языков «TurkLang-2020». (Труды конференции). Уфа: ИИЯЛ УФИТС РАН. – С. 76-80.
- [3] Abjalova M.A., Rashidova U.M. Lingvistik ta'minotni yaratishda frazeologik birliklar masalasi // "Kompyuter lingvistikasi: muammo va yechimlar". Xalqaro onlayn ilmiy-amaliy konferensiya materiallari to'plami. – Toshkent, 19-aprel, 2021-yil. – B. 22-26.
- [4] Abjalova M.A., Abdumuminov B. Lingvistik dasturlar uchun o'zbek tilidagi frazemalar bazasini yaratish tamoyillari // Kompyuter lingvistikasi: muammolar, yechim, istiqbollar / Respublika ilmiy-texnik konferensiya to'plami. Elektron nashr / ebook. – Toshkent: ToshDO'TAU, 2021. – B. 63-67.
- [5] Amanova, Lola (2014). Türkiye Türkçesindeki Deyim ve Atasözlerin Özbek Türkçesine Aktarma Problemi. TSA / YIL: 18 S: 3, Aralık, sayfa 167-177.
- [6] Rahmatullayev, Shavkat (2001). O'zbek tilining frazeologik lug'ati. – Toshkent: O'qituvchi. – 407 b.
- [7] O'zbek tilining izohli lug'ati: 80 000 dan ortiq so'z va so'z birikmasi (A. Madvaliyev tahriri ostida). 5 jildli. – Toshkent: O'zbekiston milliy ensiklopediyasi, 2006.
- [8] Turkcha lug'at
- [9] Rashidova, Umida (2018). O'zbek tilidagi somatik iboralarning semantik-pragmatik tahlili (ko'z, qo'l va yurak komponentli iboralar misolida): Filol.fan.bo'yicha falsafa doktori (PhD)... diss. – Samarqand, 2018. – 52 b.
- [10] Isayev
- [11] Uğurlu, Mustafa (2004). "Türk Lehçeleri Arasında Kelime Eş Değerliliği". *Bilig*, Sayı 29, sayfa 19-30.
- [12] Abjalova M.A., Rashidova U.M. Til korpuslarida frazemalar bazasini yaratish omillari // "O'zbek tilining milliy korpusi: muammolar va vazifalar" xalqaro ilmiy-amaliy anjuman. – Samarqand, 25.03.2023. – B. 50-54.

## EXTENDED SUMMARY

It is known that in today's globalized era, language corpora are a kind of teaching tool, a multifunctional linguistic electronic system. The presence of lexicographic resources in it along with the text base as a source of information gives users the opportunity to acquire basic knowledge and replenish their vocabulary. For this purpose, it is considered important to have a phraseme database in the lexicographic database. This article

discusses the reasons for ensuring the existence of a database of expressions in linguistic corpora and its meaning.

As a result of the analysis of idioms in Uzbek and Turkish, the following conclusions were reached:

1. Perfect fit. Idioms in typologically related Uzbek ↔ Turkish can be translated literally. For example, *eshikdan quvsang teshikdan kiradi* – if you chase him from the door, he enters through the chimney, *joni halqumiga kelmoq* – to come from the throat, *ko'z quloq bo'lmoq* – to keep an eye on, *tindirmaslik* – not to let anyone breathe, *jonini bermoq* – to spare his life. However, this fit is a small number (in green in Table 5).

2. One-to-one fit, but pragmatic application is different. Idioms translated literally can be used in special situations by representatives of both peoples. Of course, it is linked to the linguistic-cultural phenomenon. For example, although the literal translation of the idiom to endure in Uzbek is the idiom “*ko'krak kermoq*”, depending on the usage situation, it is pragmatically consistent with the idiom *tishini-tishiga qo'ymoq*.

3. Idioms that are semantically consistent but whose lexical units are inconsistent. There are idioms formed from completely different lexical units, but their content and usage situations are compatible with each other, because such idioms emerged as a result of certain situations, similes and comparisons in the life and culture of the Uzbek and Turkish peoples. For example, *tovbasiga tayanmoq* – to beat one's knee (*juda pushaymon bo'lmoq*), *ikki oyoqni anetikka tiqmoq* – to lean on the egg door (*an ishga shoshilayapti, ammo vaqt kam qolganida qo'llaniladi*), *o'ziga choh qazimoq* – to cut the branch on which it sits, *boshi ko'kka yetmoq* – to ring the skirts (*o'ta quvonmoq*).

The methods of working with phrases in the Uzbek language educational corpus were also shown. In this article, fixed combinations of words in idioms, semantic integrity and linguistic-cultural significance of idioms, and free syntactic connection are examined. Due to these properties, the importance of creating a parallel database of such stable compounds has been proven by examples. The reasons why idioms are considered as words/sentences in translation and the concept of idiom=word/sentence are explained. In general, the process of paradigmatic and syntagmatic formation of DBs acquires social significance and is considered a complex structural process, its features are more clearly revealed in the case of text and speech. However, it has been deemed appropriate to consider idioms as the structural unit of a phrase and sentence, as a ready-made word unit, and to determine their place in the form of syntactic relations.